Weakly Supervised Learning in Computer Vision Tutorial

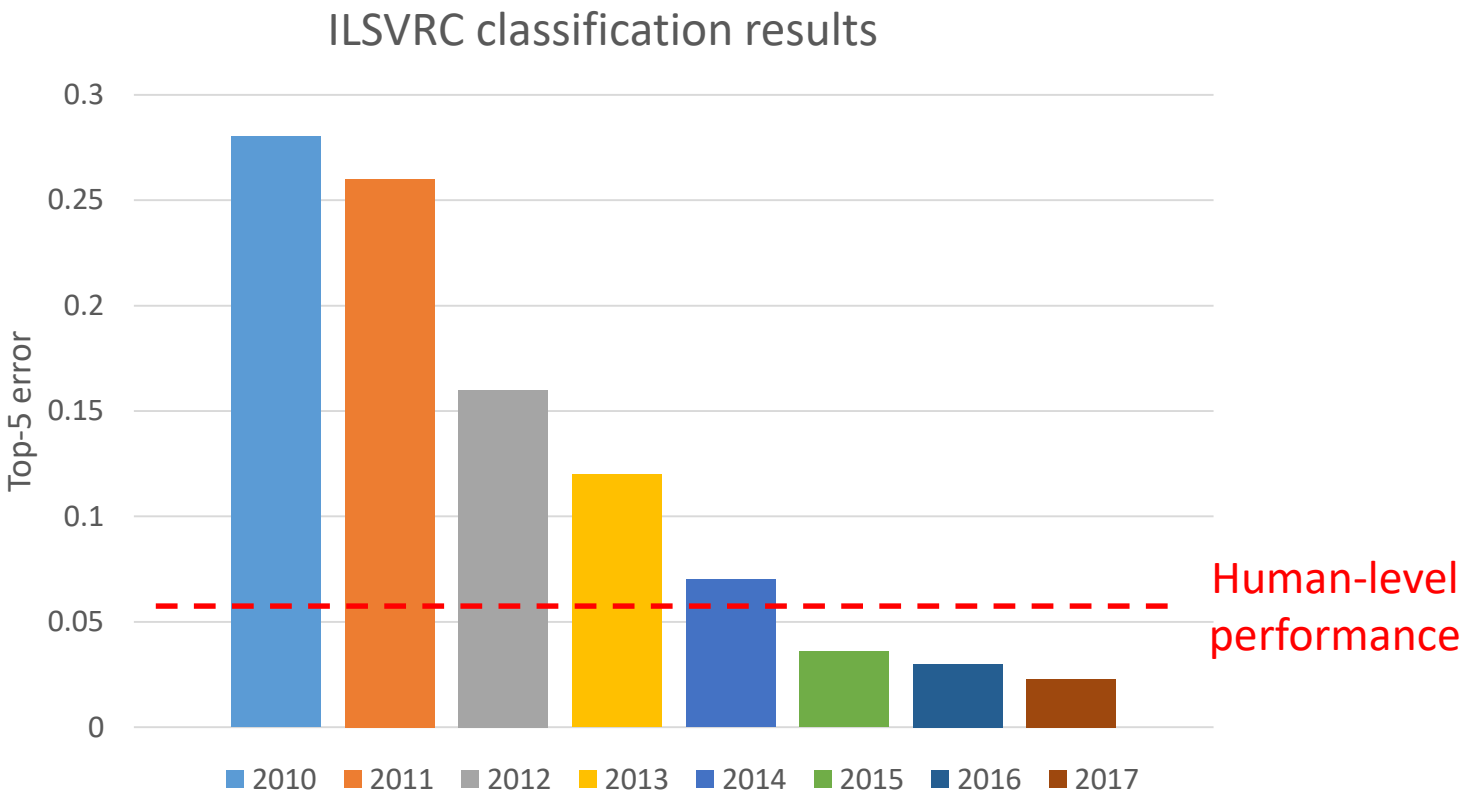
ECCV 2020

Rodrigo Benenson
Google Zurich

Seong Joon Oh
Naver

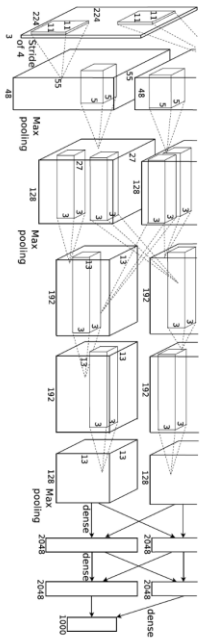
Hakan Bilen
University of Edinburgh

Human-level performance in image classification

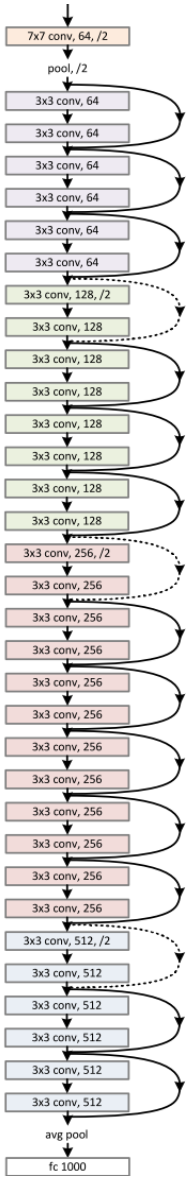


Driving forces

AlexNet



ResNet



Deep networks

Large-scale labeled dataset

Large-scale datasets

IMAGENET

1M boxes for >3000 categories

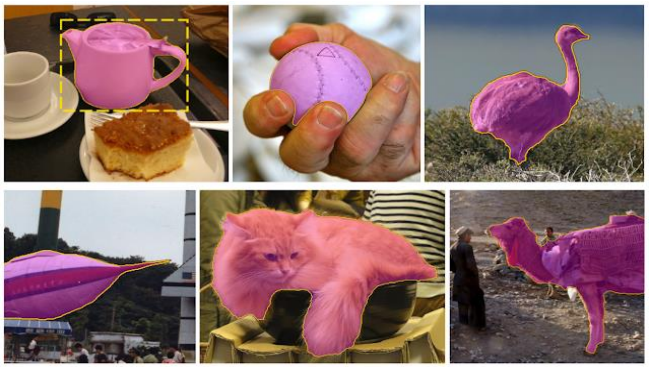
[Russakovsky IJCV 15]

COCO Common Objects in Context

118K pixelwise labeled images for 80 object and 91 stuff categories

[Lin ECCV 14, Ceasar CVPR 18]

Open Images



15M boxes on 600 categories
2,8M instance segmentations for 350 categories

[Kuznetsova IJCV 18]

Cityscapes

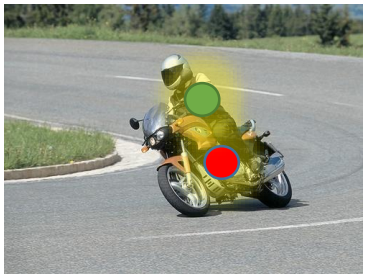
5K images for 30 categories

[Cordts CVPR 16]

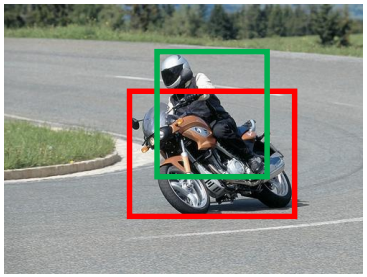
Slide credit: Jasper Uijlings

Manual supervision for object recognition

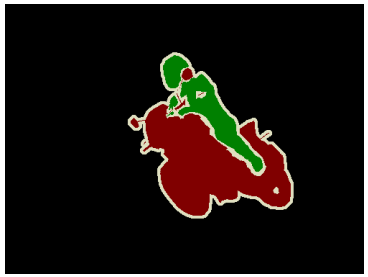
{motorbike,person}



{motorbike (point), person (point)}



{motorbike (b-box), person (b-box)}



{motorbike (pixel labels), person (pixel labels)}

1 sec per class

2.4 sec per instance

10 sec per instance

78 sec per instance

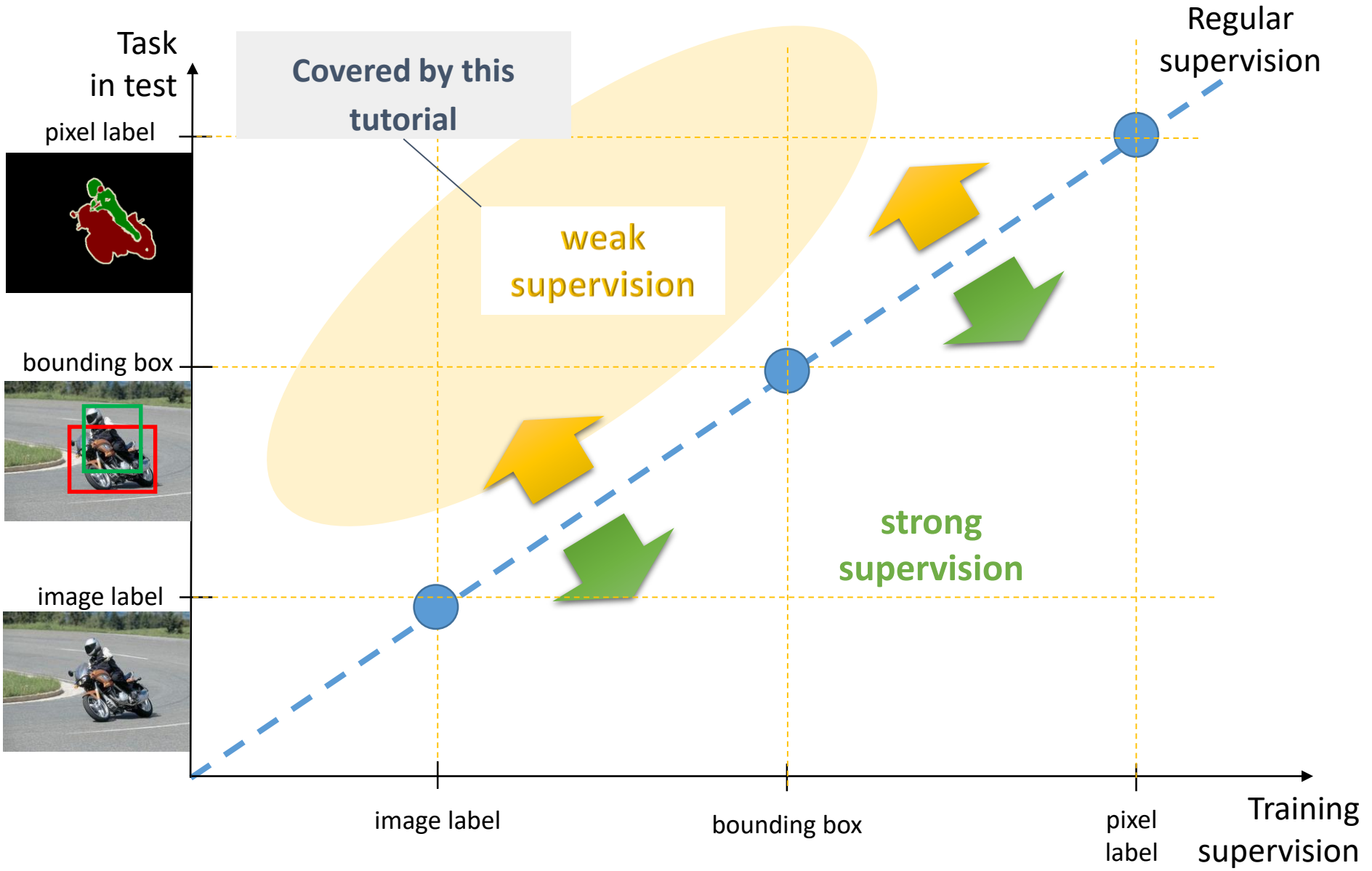
annotation time

Berman et al., What's the Point: Semantic Segmentation with Point Supervision, ECCV 16

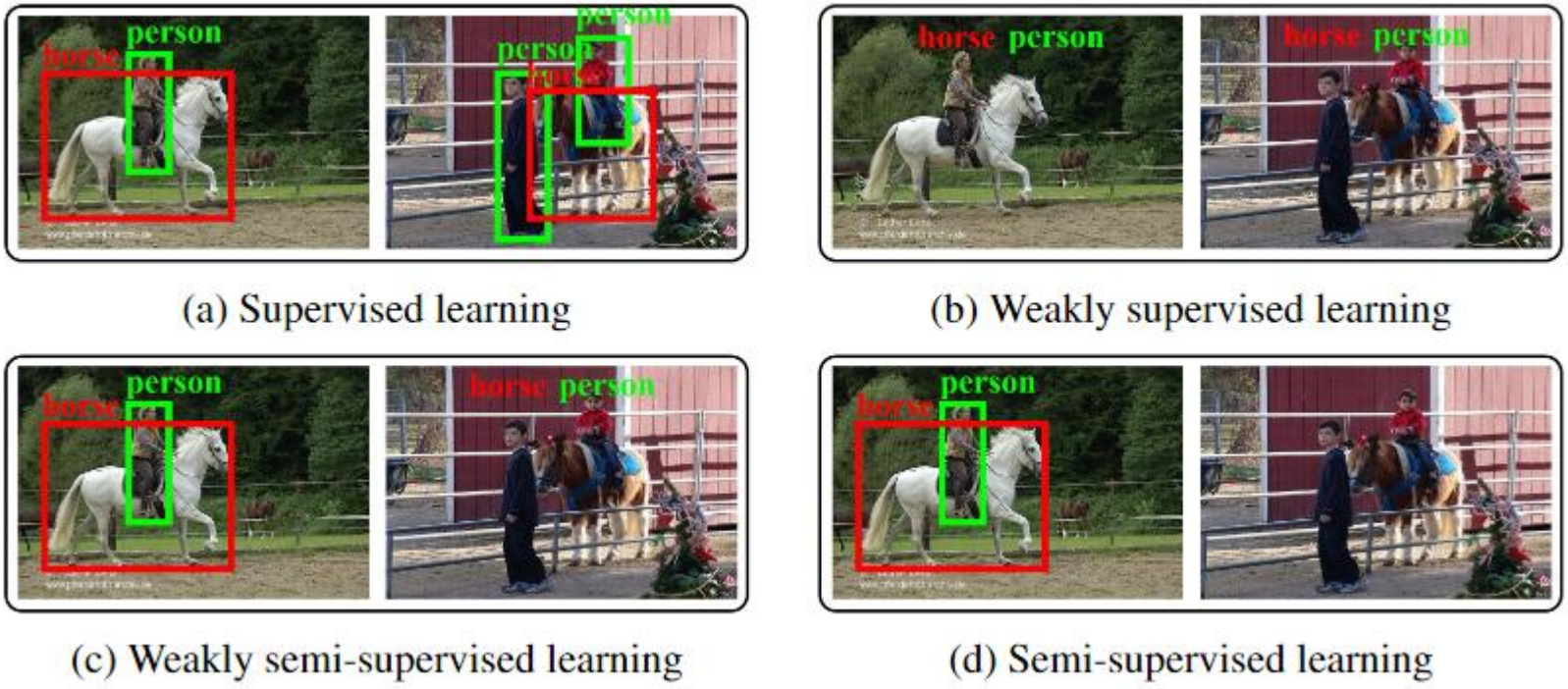
Weak supervision

Lower degree (or cheaper) annotation at train time than the required output at test time

Manual supervision for object recognition



Related topics: semi-supervised learning



(a) Supervised learning

(b) Weakly supervised learning

(c) Weakly semi-supervised learning

(d) Semi-supervised learning

- **Semi-supervised**: learns from a set of labeled (bounding box) and unlabeled images

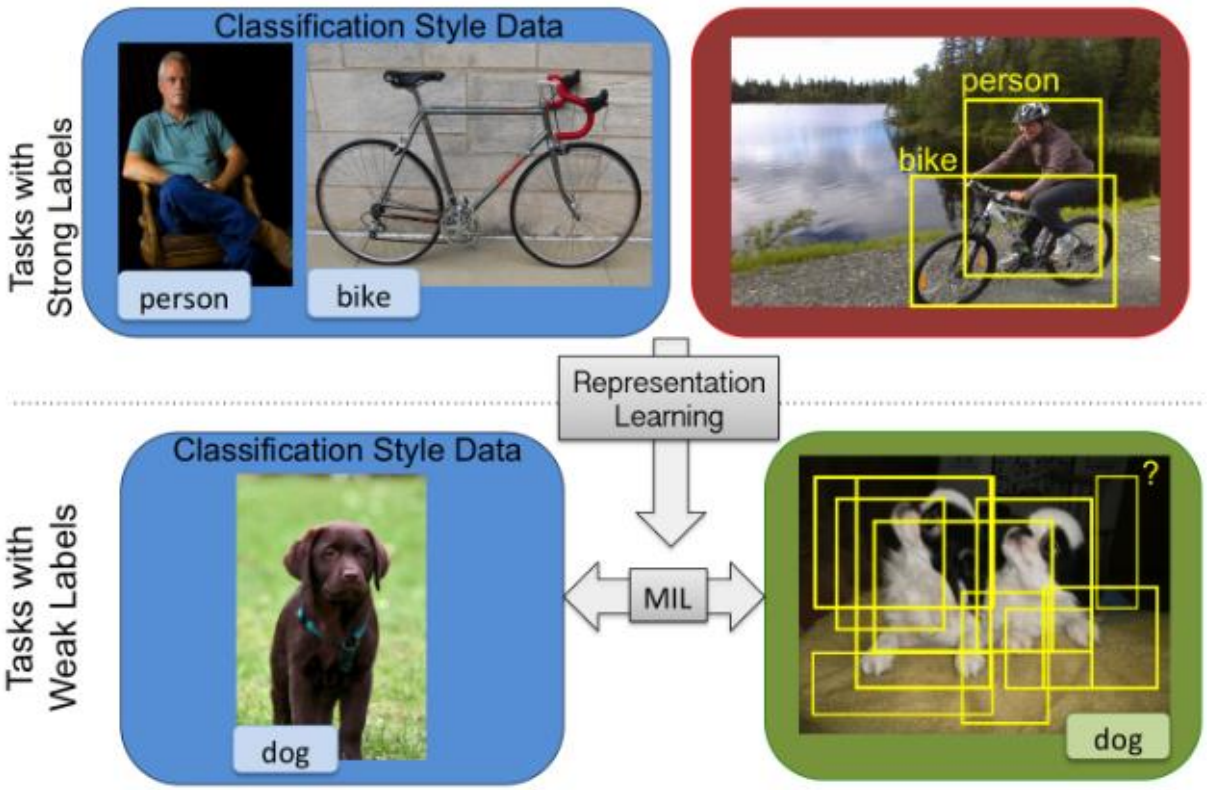
  [Wang CVPR 19]

- **Weakly semi-supervised**: learns from a set of labeled (bounding box) and weakly labeled (image label) images

  [Yan arXiv 1702.08740]

Image credit: [Jeong 19 NeurIPS]

Related topics: transfer learning



- Learns detectors for auxiliary categories with regular supervision (bounding box)
- Transfers the knowledge from auxiliary categories to weakly supervised categories

[Hoffman NIPS 14], [Tang CVPR 16]

Image credit: [Hoffman NIPS 14]

Related topics: Self-supervised learning

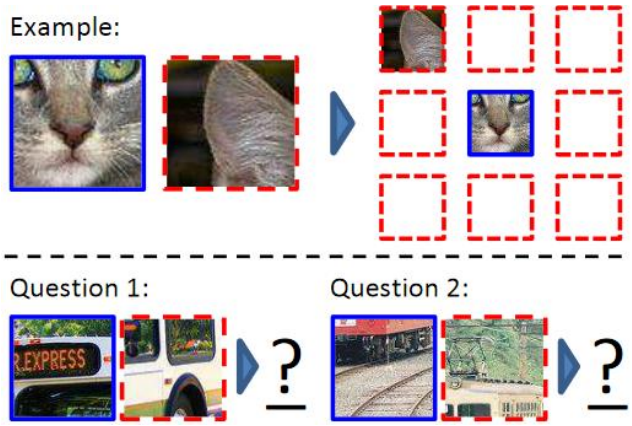


Image credits: [Doersch, ICCV 15]

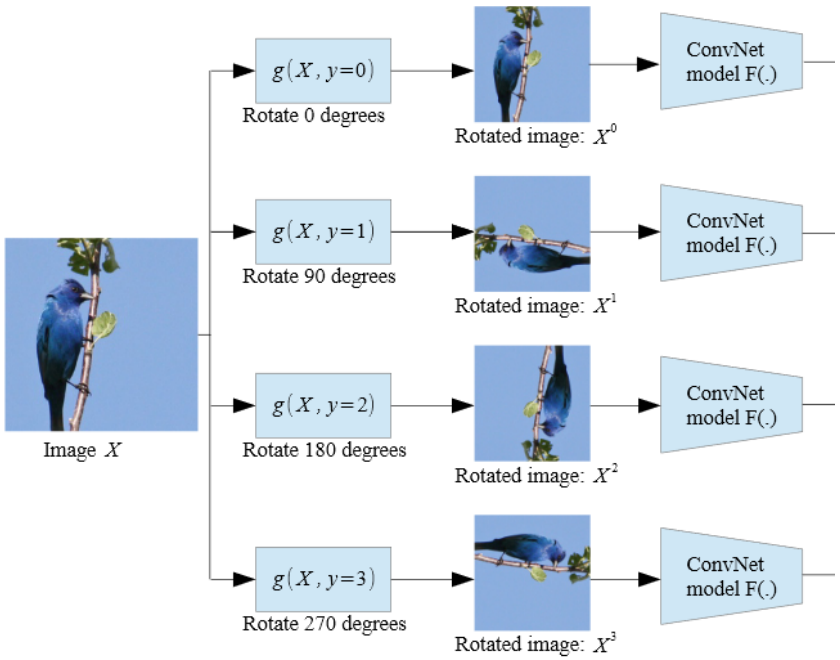


Image credits: [Gidaris ICLR 18]

<u>Self-supervised learning</u>

- Defines a pretext task and learns representations in an unsupervised way
- Transfers pre-trained weights to a down-stream task
- Transfer-learning requires regular annotations

<u>Weakly supervised learning</u>

- Typically uses pretrained networks and fine-tunes them by using weak supervision

Both problems reduce dependency to annotations and are complementary!

Generic Object Recognition

Intra-class variations

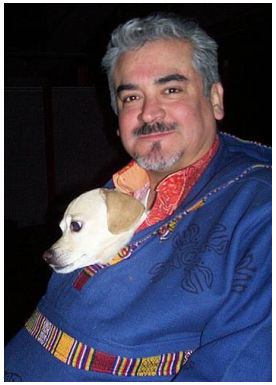
- Appearance
- Viewpoint
- Scale
- Aspect ratio

Background clutter

Occlusions

Weakly Supervised Learning

Parts vs object



Question: What is a person?

a) Face

b) Face + upper body

c) Face + whole body ☺

Weakly Supervised Learning

Context vs object

Question: What is a train?

a) Rail

b) Train + Rail

c) Train ☺

Weakly Supervised Learning

Multiple vs single instance

Question: What is a sheep?

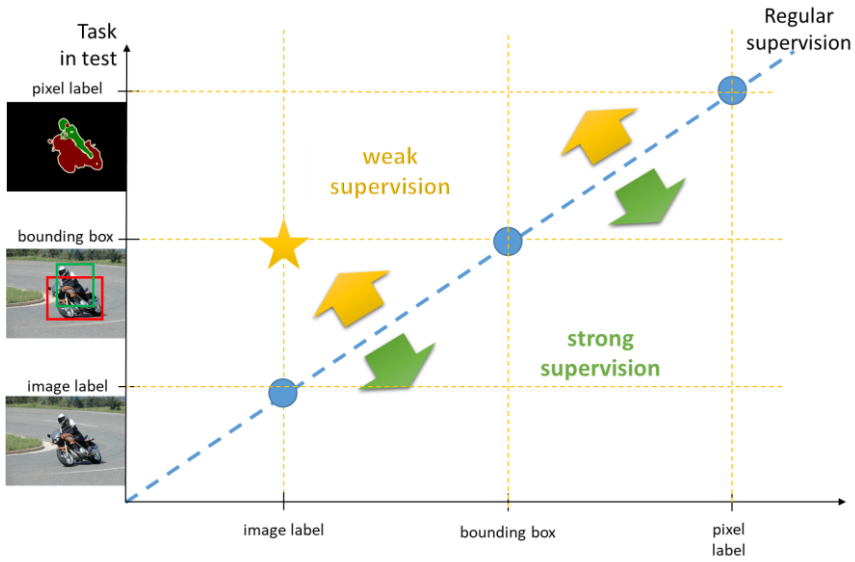
a) 2 sheeps

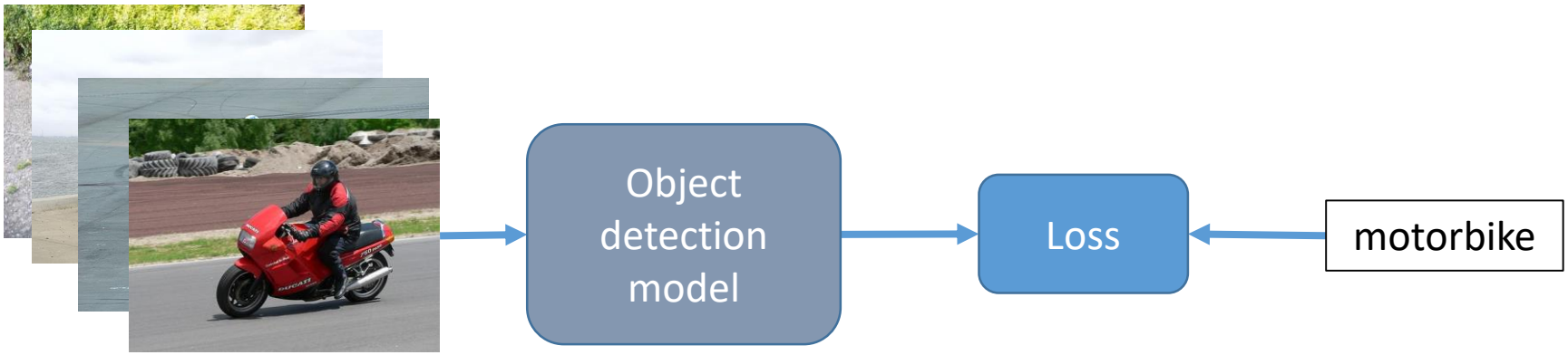
b) 3 sheeps + grass

c) 1 sheep ☺

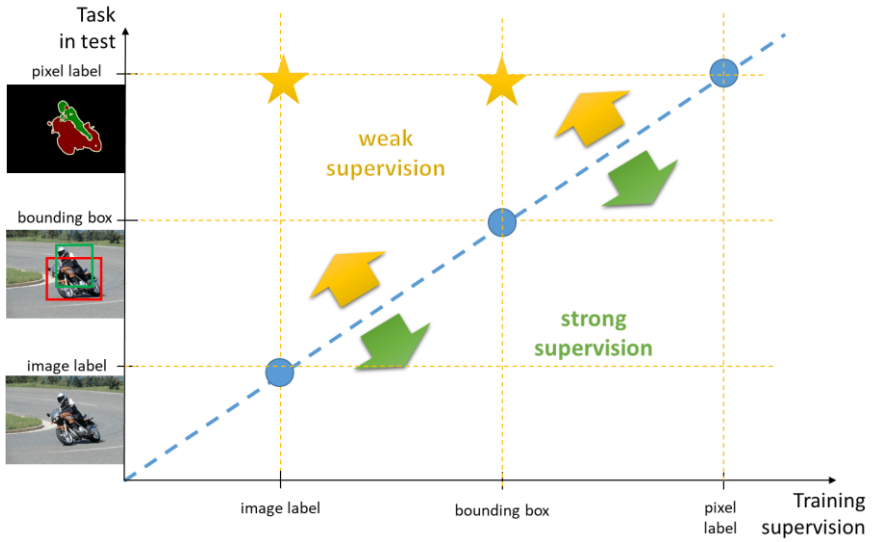
Hakan Bilen

- Focuses on object detection with image-level supervision
- Review ideas from 50+ papers
- Provides a framework to the problem
- Discuss challenges and existing solutions





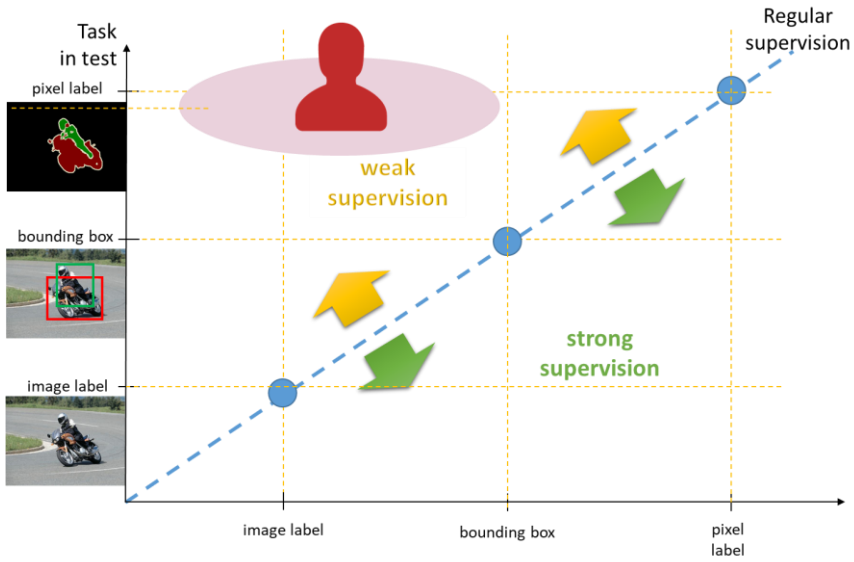
Rodrigo Benenson



- Shares the insights of 40+ papers.
- Provides a framework to read (and write) weakly supervised papers.
- Discusses use of image-labels and box-labels, for semantic and instance segmentation.

Rodrigo Benenson

- Discusses the gap between papers and practice.
- Describes active learning strategies for box and segmentation annotations.
- Presents couple of example papers in the field.

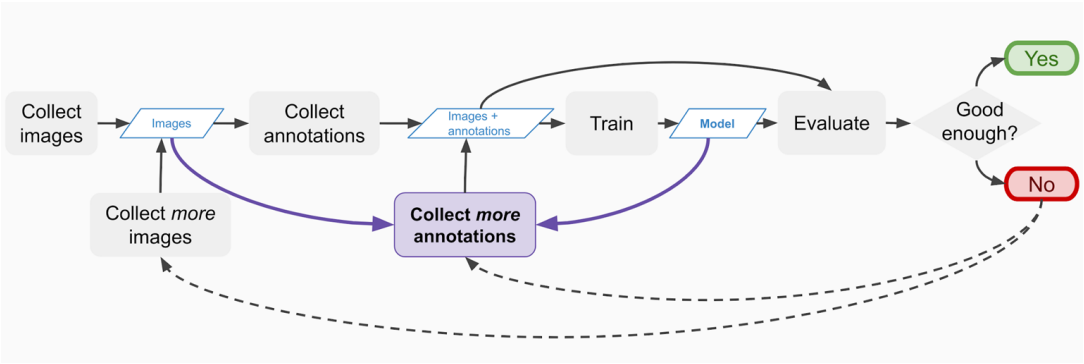


Detection

Which boxes to add?

Semantic labeling

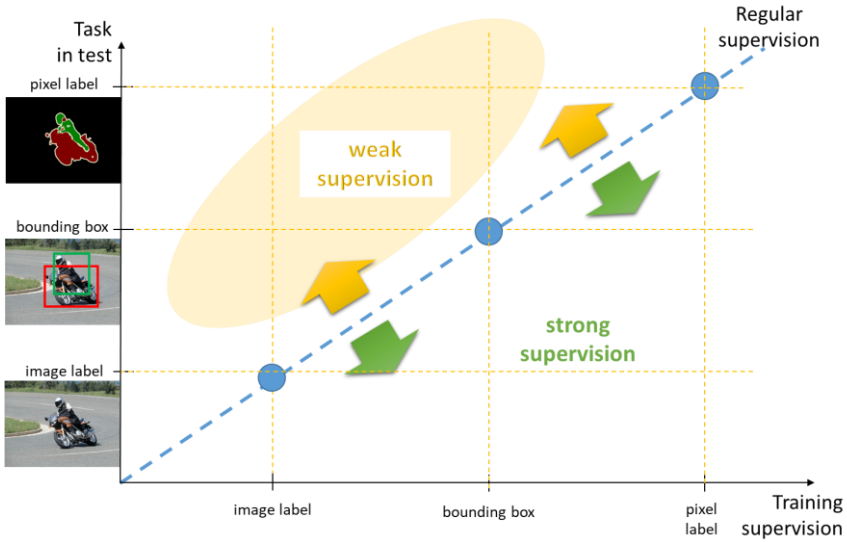
Which pixels to add?

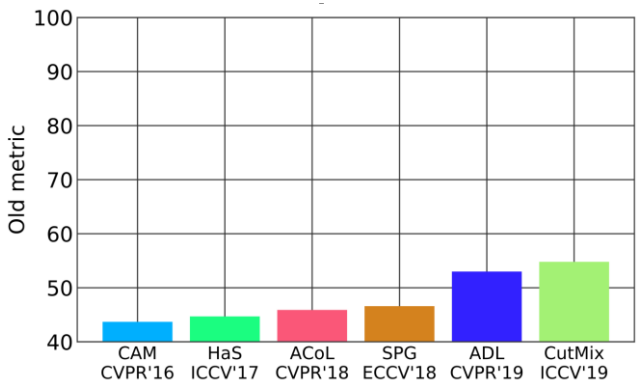


Evaluating weakly-supervised models

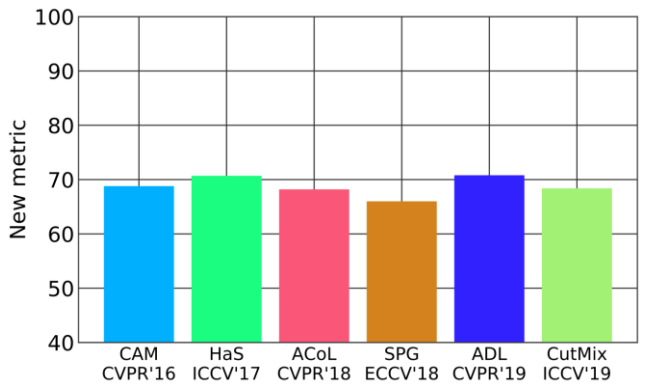
Seong Joon Oh

- Discusses evaluation problem in weakly-supervised learning
- Presents a critical view on the current evaluation practices
- Provides fair evaluation methodologies
- Re-evaluate existing methods in the new benchmark





Improvement according to papers 2016 - 2019



No real improvement under the corrected evaluation.

Please visit our Q/A sessions

All the videos and slides will be available in our tutorial webpage:

https://hbilen.github.io/wsl-eccv20.github.io/